1966

Study of the relationship between item ambiguity susceptibility to social influence and the double agreement phenomenon

Loren Kent McBride
The University of Montana

A STUDY OF THE RELATIONSHIP BETWEEN ITEM AMBIGUITY, SUSCEPTIBILITY TO SOCIAL INFLUENCE AND THE DOUBLE AGREEMENT PHENOMENON

By

Loren Kent McBride

B. S. University of Idaho, 1964

Presented in partial fulfillment of the requirements for the degree of

Master of Arts

UNIVERSITY OF MONTANA

1966

Approved by:

Gary P. Moran

Chairman, Board of Examiners

Dean, Graduate School

JUN 6 1966

Date

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 - 1346

ACKNOWLEDGMENT

I am deeply indebted to Dr. Raymond M. Gold, Mr. Alan Klockars, Committee members, and in particular to Dr. Gary P. Moran, Committee Chairman, under whose expert guidance and friendly assistance this study has been pursued.

Loren K. McBride

Introduction

The so-called "double agreement phenomenon" is a topic of current interest in the social psychological literature. "Double agreement" refers to a subject's endorsement of an item as well as its subsequent reversal.

Peabody (1961) infers from the occurrence of double agreement to items (and their reversals) on the California F test (authoritarianism), Anti-Semitism scale (Adorno, 1950), and Dogmatism (Rokeach, 1956) scale a so-called "agreement response set." (ARS) That is, the content of the items is presumed to be irrelevant. This view, of course, would logically lead to a tentative rejection of the so-called authoritarianism syndrome, as defined by F, Dogmatism (D) or Anti-Semitism (A-S) content. Sources relevant to these scales include Christie and Cook (1958), and Christie, Havel and Seidenberg (1958).

Peabody further suggested that agreement response set as a determinant of item endorsement does not operate in a blanket fashion. Subjects do not endorse items indiscriminately; only those items whose content are "ambiguous" are subject to double agreement. Peabody did not further specify the concept "ambiguous," nor was empirical support for this ad hoc hypothesis presented.

Rokeach (1963) discussed Peabody's "ambiguity" hypothesis in detail and concluded that no experimental demonstration of its validity has thus far been presented. Unfortunately, he also neglected to present data relevant to this hypothesis.

Peabody (1966) in a recent paper reaffirmed his original contention that item ambiguity predicts double agreement. Oddly, he neither reviewed or presented data to substantiate his position. Such data is apparently not at hand.

The main objective of this study is an identification of conditions under which double agreement occurs. It is proposed that double agreement bears a positive relationship to item ambiguity. A criterion of ambiguity will be presented. This represents a test of Peabody's hypothesis, for which no empirical evidence appears to be available.

Secondly, it is proposed that individual differences in predisposition to acquiesce can be identified. Subjects characterized by susceptibility to social influence will show more acquiescence. Specifically, it is proposed that high need for approval subjects will show more double agreement than do low need for approval subjects. The approval motive construct's utility for predicting many behavioral phenomena is established (Crowne and Marlowe, 1964). It is implicitly assumed that approval-dependent subjects feel they are behaving in a socially desired manner by complying with demand implications of the experiment, that is, by agreeing to most items presented them by E.

Recapitulating, the following experimental hypotheses will be tested:

Hypothesis 1: High need for approval subjects show more double agreement than do low need for approval subjects.

Hypothesis 2: The greater the ambiguity of combined original and reversed item, the higher the incidence of double agreement.

Procedure

Materials

Materials utilized in this investigation included items drawn from original F, Dogmatism, and Anti-Semitism scales as used by Peabody (1961), Peabody's reversals of these items, ambiguity rating sheets with appropriate instructions, and copies of the Marlowe-Crowne Social Desirability Scale (Crowne and Marlowe, 1964, pp. 23-24). Original and reversed scales appeared separately in test booklets in one of three counterbalanced orders. Items were numbered consecutively so that subjects were confronted with a single, eighty-eight item questionnaire. In addition, Peabody's reversals were placed in individual scales in an order reversing their sequential appearance in original scales. This procedure was employed to lessen the probability that a subject would remember the exact content of original items. For the rating session, statements were separated into original and reversed booklets, but were neither counterbalanced nor disguised so as to appear as a single set of statement.

Subjects

One hundred eighty freshmen students participated as subjects in the actual testing phases and were randomly assigned to separate subgroups. This group was composed of one hundred eight males and seventy-two females, with a mean age of 19.6.

Ambiguity Ratings

The operational definition of "ambiguity" involves a dimension which may be called, "the clarity of the statement." Subjects participating in this part of the study constituted a sample of judges who rated the ambiguity of items on the basis of this definition. Original and reversed items were placed into separate booklets and randomly distributed to these subjects who were then given the following instructions:

> You are to assess the ambiguity of the following statements in terms of the degree to which you understand their content and meaning. We are not concerned with whether you do or do not agree with the item; nor are we concerned with whether you would admit that it is true of you. We are solely concerned with whether or not you are clear as to what is being asked. Please rate each item's ambiguity by placing a mark at the appropriate point on the scale for each item. Two poles are presented on each scale: "Clear, straightforward," and "ambiguous." You are to judge where on this continum each item falls. Be sure and mark each statement; do not leave any out.

Eight-point scales were utilized for the ambiguity ratings. Ambiguity values for each original and reversed item were determined by finding their mean value as rated by the subjects. The ambiguity of item-pairs was calculated from these data. Two independent groups rated the items. Group A rated originals, group B reversals.

Testing Phase

Subjects who participated in testing phases were given the following directions, adapted from Peabody (1961):

> PUBLIC OPINION SURVEY: This is a study of what the general public thinks about a number of social questions. The best answer to each statement below is your personal opinion. We have tried to cover many different points of view. You may find yourself agreeing with some of the statements strongly, disagreeing just as strongly with others, and perhaps feeling less strongly about others. Whether you agree or disagree with any statement, you can be sure that many people feel the same way you do.

Do not place answers or other marks on the questionnaire. If you agree with the statement, place a line in alternative one on the IBM sheet. If you disagree with the statement, darken alternative two. Please be sure you have answered all of the questions and have your answers in the right places on the IBM sheet.

Subjects who volunteered for the afternoon session were assigned to group A; those who signed up for the evening session were assigned to group B. Group A subjects first answered the Marlowe-Crowne Social Desirability Scale, then the original authoritarianism scales. At the end of this session, they were instructed to return at the same hour for another experiment but were told nothing else. Two days later the same subjects answered the reversed scales. Group B answered the Marlowe-Crowne test and then participated in the opposite order, that is, these subjects answered reversals at the first sessions, original scales two days later.

Results

Data used to test Hypothesis 1 appear in Table 1.

TABLE 1

Correlations between Marlowe-Crowne Scores and Double Agreement Scores on F, D, and A-S Scales, and Summated over Scales.

Scale	Pearson r	P*	Mean DA Score	Range of DA Scores
F	.36	.001	6.79	1-17
D	.44	.001	10.00	3-36
A-S	.25	.001	2.44	0-15
Over scales	.44	.001	19.23	4-48

P* values are one-tailed

Mean values listed refer to the average of subjects' double agreement scores (DA) obtained on the F, Dogmatism, and Anti-Semitism scales, as well as to total double agreement scores summated over all three scales. These values are 6.79, 10.00, 2.44, and 19.23 respectively. Ranges for these measures are also provided.

For providing a specific test of the validity of the hypothesis that need for approval bears a positive linear relationship to double agreement, Pearson rs were computed not only between subjects' Marlowe-Crowne and double agreement scores on the F, Dogmatism, and Anti-Semitism scales separately, but also for double agreement scores summated over all scales. As shown in Table 1, the values of r associated with these measures are .36 for the F scale, .44 for the Dogmatism scale, .25 for the Anti-Semitism scale, and .44 over scales. With df equal to 178 in all cases, all correlations are significant beyond the

.001 level. These results clearly indicate a strong relationship between need for approval and acquiescence. Hypothesis 1, under conditions described in the present sample, must be considered tenable.

Ambiguity ratings for individual original and reversed scales in all three tests appear in appendices 1, 2 , and 3. Values for individual original and reversed items were subsequently averaged to yield the ambiguity scale values for each of the item-pairs. Additionally, these appendices include double agreement responses given to each item, expressed both as frequencies and percentages. Item-pair ambiguity values were then correlated with the frequency of double agreement associated with the particular stimuli.

Pearson rs serve to test the validity of the hypothesis that item ambiguity bears a positive linear relationship to double agreement. Coefficients are as appear in Table 2.

TABLE 2

Correlations between Rated Ambiguity and Frequency of Double Agreement on F, D, and A-S Scales.

Scale	Pearson r	P*
F	.88	.001
D	.83	.001
A-S	.65	.001

P* values are one-tailed

The values of r are, respectively, .88 for the F scale, .83 for the Dogmatism scale, and .65 for the Anti-Semitism scale. With df equal

to 26, 38, and 18, all correlations are significant beyond the .001 level. These results indicate an extremely strong positive relationship between rated ambiguity of item-pairs and frequency of double agreement on all three measures of authoritarianism. Hypothesis 2, under the conditions described, must likewise be considered tenable.

Discussion

The major reason conducting the present investigation was to provide a direct empirical test of Peabody's ad hoc hypothesis that double agreement to original and reversed authoritarianism scales is a positive function of the scales' ambiguity. The present data clearly indicate that, within separate tests, the more ambiguous an F, Dogmatism, or Anti-Semitism scale item, the more likely will be double agreement (DA) to that item and its reversal.

To date, two major criticisms have been formulated against Peabody's assertion that the phenomenon of agreement response set (ARS) casts doubt on the ability of authoritarianism scales to measure the so-called "authoritarian personality."

Samuelson (1964) suggested that the occurrence of double agreement in Peabody's data (1961) may be a function of asymmetrical reversals of original items. Should this be the case, endorsement of both original and reversal could arise because both statements fall within a given subject's agreement continuum, and as such, double agreement constitutes consistent responding on his part. Double agreement in this case would not militate against a content analysis of endorsement of original authoritarian items. Although this interpretation appears a priori

reasonable, contrary evidence has been presented (Peabody, 1966; Miklich, 1966). Miklich's data are particularly damaging, in that a Kuder-Richardson formula (form 20) applied to original and reversed items (treated as item-pairs) produced a value of .54. Peabody's reversals are in fact relatively homogeneous.

Rokeach (1963) contended that no evidence exists to support Peabody's agreement response set hypothesis, either to the effect that authoritarianism scale items are ambiguous, or, even granting ambiguity, that it predicts to double agreement. He presents an alternative hypothesis to the effect that a given subject endorses both an original and reversed item because he is responding truthfully to the former and deliberately lying to the latter for reasons of his own. One such reason is that an authoritarian subject will view original authoritarianism items as harmless. Their reversals will be falsely agreed to because such agreement is socially desirable.

Rokeach's hypothesis has to date received no empirical verification and is exclusively supported by anecdotal accounts. Additionally, it appears untenable on two grounds. First, Stanley and Martin (1964) found a significant negative relationship between original Dogmatism scale items and the Martin Social Desirability Scale, that is, agreement to Dogmatism items is not socially desirable behavior. Further, scores on a reversed Dogmatism scale failed to correlate significantly with the Martin test for a group of high acquiescent subjects. On the other hand, the present data clearly demonstrate that frequency of double agreement is a positive function of the relative ambiguity of item-pairs. Rokeach's formulation to date appears to be without support.

Peabody's basic rationale contends that a significant proportion of high "authoritarian" scores on F, Dogmatism, and Anti-Semitism scales represent an artifact produced by an interaction between subject and content variables. Following Cronbach's (1946) well-known analysis of agreement set tendencies, he suggests that agreement response set to original statements constitutes a subject characteristic stemming from response uncertainty created when subjects are forced to deal with ambiguous items.

How ambiguous are authoritarianism items? The evidence in the literature seems to reflect little other than rank opinion. The mean ambiguity scale values for the F, Dogmatism, and Anti-Semitism items in the present sample are 3.84, 3.77, and 4.31. None of these values indicate ambiguity in an absolute sense as they all fall at or below the mid-point of an eight-point scale. Relatedly, Rokeach (1963) is "puzzled" by the high incidence of double agreement on the Anti-Semitism scale in Peabody's 1961 data. In his opinion, the Anti-Semitism scale is "known" to be less ambiguous than either F or Dogmatism scales. This study, however, indicates that the Anti-Semitism scale is the most ambiguous of the three used, significantly more so than the F ($p < .02$) or Dogmatism ($p < .01$) scales. Again, Rokeach appears in error.

Another relevant consideration concerns the frequency of double agreement. For the F, Dogmatism, and Anti-Semitism scales respectively, double agreement represents 24%, 25%, and 12% of subjects' total responses. Although the incidence of such behavior is substantial, it nevertheless is approximately one-third less than that reported by Peabody (1961). Of greater significance, however, is the fact that the

smallest frequency of double agreement occurred on the scale (Anti-Semitism) rated as most ambiguous. This finding is noteworthy in that Peabody reasoned that the more ambiguous a scale, the greater the frequency of double agreement to be expected. The present data warrant the conclusion that Peabody's hypothesis that double agreement is a function of the ambiguity of the test's items bears modification. The relative ambiguity of items in a given scale and not a test's overall ambiguity seems to relate to double agreement on that test.

An additional finding concerns the substantial relationship demonstrated between double agreement scores and subjects' need for approval scores. Confirmation of Hypothesis 1 warrants the conclusion that many high authoritarian scores may be an artifact created by an interaction between high approval need and perceived social pressures inherent in any group testing situation. This finding serves to specify one significant population variable involved in the erroneous identification of bogus "authoritarians." Given the nature of approval dependent individuals (Crowne and Marlowe, 1964), the finding increases the credibility of an explanation of "bogus" authoritarians in terms of situational factors, specifically, social pressure.

In summary, the results of this investigation indicate that Peabody's agreement response set construct is essentially correct but nevertheless requires modification. While some individual scale items are in fact highly ambiguous, the assertion that the scales themselves are ambiguous in an <u>absolute</u> sense remains unsupported. Unfortunately, little evidence is available as to the absolute level of ambiguity on other like instruments, e.g., the MMPI. Perhaps authoritarianism tests are more ambiguous

than similar diagnostic tools. The issue invites investigation rather than additional pontification. Finally, the present data indicate that need for approval is a subject characteristic worthy of consideration in analyzing the double agreement problem. Data provide concrete support for Peabody's (1961, 1966) contention that agreement set response does not operate in blanket fashion.

High authoritarian scores may in many cases represent little more than the result of ambiguous measuring scales and perceived social pressures inherent in the testing situation itself. There appears little need to invoke a content syndrome. Accurate identification of "authoritarians" by means of the content of traditional scales must be viewed with profound skepticism. We feel that Samuelson, Rokeach, and others who would preserve the credibility of these scales must demonstrate their construct validity. Only a methodology that goes beyond traditional construct validation appears capable of satisfying necessary criteria. Constructive steps toward developing such procedures appear in Loevinger (1957) and Campbell and Fiske (1959). The essential requirement of both models is that construct validity, even if established by conventional procedures, must in addition include a demonstration that the measure under consideration correlate neither with known or suspected error sources, nor with secular trends, such as agreement response set, social desirability, or examination variables.

These additional criteria for establishing construct validation seem essential in consideration of the authoritarian syndrome precisely because authoritarianism measures <u>do</u> meet traditional and less rigorous requirements of correlation with postulated aspects of the syndrome.

Regardless, as we have clearly seen, the construct validity of these measures is seriously in question. By endorsing more rigorous requirements for validation of measures of authoritarianism, we may yet identify some "content" which accurately characterizes authoritarians. This seems a worthwhile goal and one of greater ultimate significance than the tedious, if essential, demonstrations that the present postulated content of authoritarian scales finds scant support in the data and is attributable principally to both sources of error variance and secular trends.

References

Adorno, T., Frenkel-Brunswik, Else, Levinson, D., and Sanford R. The Authoritarian Personality. New York: Harper, 1950.

Campbell, D., and Fiske, D. "Convergent and discriminant validation by the multitrait-multimethod matrix," Psychological Bulletin, 1959, 56, 81-105.

Cristie, R., and Cook, Pauline. "A guide to published literature relating to the authoritarian personality," Journal of Psychology, 1958, 45, 171-200.

Cristie, R., Havel, Joan, and Seidenberg, B. "Is the F scale irreversable?" Journal of Abnormal and Social Psychology, 1958, 56, 143-149.

Cronbach, L. "Response set and test validity," Educational and Psychological Measurement, 1946, 6 475-494.

Crowne, D., and Marlow, D. The Approval Motive: Studies in Behavioral Dependence. New York: John Wiley and Sons, Inc., 1964.

Loevinger, Jane. "Objective tests as instruments of psychological theory," Psychological Reports, 1957, 3, 635-694.

Miklich, D. "Peabody's agreement response set measure: Reply to Samuelson," Psychological Reports, 1966, 18(1), 200-202.

Peabody, D. "Attitude content and agreement sets in scales of authoritarianism, dogmatism, anti-Semitism, and economic conservatism," Journal of Abnormal and Social Psychology, 1961, 63, 1-11.

Peabody, D. "Authoritarianism scales and response bias," Psychological Bulletin, 1966, 65(1), 11-23.

Rokeach, M. "Political and religious dogmatism: an alternative to the authoritarian personality," Psychological Monographs, 1956, No. 18 (Whole No. 425).

Rokeach, M. "The double agreement Phenomenon: Three hypotheses," Psychological Review, 1963, 67, 304-309.

Samuelson, F. "Agreement set and anticontent attitudes in the F scale: a reinterpretation," Journal of Abnormal and Social Psychology, 1964, 68, 338-342.

Stanley, G., and Martin, J. "How Sincere is the Dogmatist?" Psychological Review, 1964, 71, 331–334.

Appendix 1

Mean Ambiguity Values and Double Agreement Responses: F Scale

Item	Original	Reversed	Item Pair	Freq. DA	Per cent DA
1.	2.20	2.40	2.30	12	6.66%
2.	4.14	4.89	4.64	78	43.33%
3.	3.86	2.64	3.25	33	18.33%
4.	4.10	4.83	4.46	64	35.55%
5.	4.40	4.51	4.45	50	27.77%
6.	3.96	3.08	3.52	40	22.22%
7.	3.48	2.18	2.83	13	7.22%
8.	3.33	2.89	3.11	17	9.44%
9.	3.69	3.82	3.75	32	17.77%
10.	5.09	2.78	3.93	43	23.88%
11.	2.95	2.69	2.82	13	7.22%
12.	3.39	3.39	3.39	28	15.55%
13.	2.77	4.34	3.55	29	16.11%
14.	3.78	5.14	4.46	41	22.77%
15.	4.35	3.13	3.74	31	17.22%
16.	3.41	3.12	3.26	15	8.33%
17.	3.55	4.69	4.12	39	21.66%
18.	3.57	3.87	3.72	37	20.55%
19.	3.35	3.54	3.44	20	11.11%
20.	3.57	4.98	4.27	37	20.55%
21.	4.84	5.59	5.21	92	51.11%
22.	3.29	3.48	3.38	33	18.33%
23.	3.48	3.57	3.52	33	18.33%
24.	5.50	3.89	4.69	87	48.33%
25.	3.70	4.43	4.07	33	18.33%
26.	4.47	4.08	4.27	49	27.22%
27.	4.39	4.11	4.25	58	32.22%
28.	4.40	3.86	4.15	37	20.55%

Appendix 2

Mean Ambiguity Values and Double Agreement Responses: D Scale.

Item	Original	Reversed	Item Pair	Freq. DA	Per cent DA
1.	4.31	3.77	4.04	46	25.55%
2.	4.61	4.21	4.41	66	33.33%
3.	4.41	4.54	4.47	69	38.33%
4.	4.40	1.81	3.10	32	17.77%
5.	4.52	4.31	4.41	47	26.11%
6.	2.61	4.93	3.77	33	18.33%
7.	2.71	3.49	3.10	37	20.55%
8.	3.76	5.78	4.77	119	66.22%
9.	4.34	4.33	4.33	60	33.33%
10.	3.35	3.77	3.65	52	28.33%
11.	4.67	4.50	4.58	78	43.33%
12.	4.07	3.07	3.57	33	18.33%
13.	3.98	2.11	3.04	20	11.11%
14.	4.31	3.13	3.72	46	25.55%
15.	3.53	3.42	3.47	63	35.00%
16.	3.49	2.36	2.92	20	11.11%
17.	4.06	3.83	3.94	32	17.77%
18.	3.39	5.34	4.36	57	31.66%
19.	4.44	3.98	4.21	53	29.44%
20.	4.82	3.07	3.95	38	21.11%
21.	4.17	4.17	4.17	48	26.66%
22.	3.76	5.43	4.59	97	53.88%
23.	3.83	2.55	3.19	19	10.55%
24.	3.60	3.35	3.47	24	13.33%
25.	2.72	4.33	3.52	24	13.33%
26.	3.95	4.17	4.06	77	42.77%
27.	3.86	4.24	4.05	24	13.33%
28.	4.07	4.47	4.27	51	28.33%
29.	4.40	3.66	4.03	38	21.11%
30.	3.74	3.01	3.37	23	12.88%
31.	3.04	2.05	2.54	7	3.88%
32.	3.56	3.56	3.56	36	20.00%
33.	3.25	4.37	3.37	38	21.11%
34.	3.06	3.69	3.37	35	19.44%
35.	2.58	2.52	2.55	6	3.33%
36.	2.47	2.87	2.67	17	9.44%
37.	3.52	3.34	3.43	35	19.44%
38.	3.20	4.52	3.86	41	22.88%
39.	3.93	5.41	4.67	76	42.22%
40.	3.54	5.51	4.52	81	45.00%

Appendix 3

Mean Ambiguity Values and Double Agreement Responses: A-S Scale.

Item	Original	Reversed	Item Pair	Freq. DA	Per cent DA
1.	4.79	4.39	4.57	55	30.55%
2.	3.13	2.82	2.97	4	2.22%
3.	5.30	3.10	4.20	16	8.88%
4.	5.42	2.70	4.06	11	6.11%
5.	4.60	3.13	3.86	8	4.44%
6.	5.19	2.74	3.96	9	5.00%
7.	4.37	3.07	3.72	13	7.22%
8.	4.24	3.64	3.94	26	14.44%
9.	4.60	2.14	3.37	8	4.44%
10.	3.83	4.02	3.92	27	15.00%
11.	3.88	2.54	3.19	8	10.55%
12.	4.95	3.58	4.26	12	6.66%
13.	4.10	3.16	3.63	13	7.22%
14.	4.35	4.44	4.39	47	26.16%
15.	3.22	3.10	3.16	13	7.22%
16.	3.85	4.60	4.22	38	21.11%
17.	3.93	1.83	2.88	9	5.00%
18.	4.01	3.74	3.87	21	11.66%
19.	4.41	3.53	3.96	21	11.66%
20.	4.10	3.65	3.87	23	12.88%